

Figure 1

Anopheles gambiae arrestin 1 cDNA sequence (SEQ ID NO: 1)

5 ACAGGAACGACGGTTGTGATCCCTCCACTGGTGGTGACACGAATCATAAGCATTATTCATAACCT
AAAAAAACAAAATCTACAAAAAAAGCTTCATTCCCATCGAAAAAAACTTCTTGTGAAATCAACCG
AGCTAACAAACAAACATCCTGTGAAAATCTAGCAGTGAAAGTGTGATATCGTATACCTGTACCTG
TAAACCGTTGTGCGCGTGTGCCTTGTGTATCAATTGTGGAAAACAGAAAATACATCAAAA
10 **TGGTTACAATTCAAAGTCTTCAAGAAGTGC**GGCCCTAATGGAAAGGTTACGCTGTACATGGG
CAAGCGTGACTTGTAGACCACGTTCCGGCGTTGAACCGATCGATGGTATCGTCGTCCTCGAT
GATGAGTACATTCGTGACAACCGTAAGGTATTGGTCAAGCTGCAGTTCCCGCTACGCC
GCGAAGAGGACGAGGTGATGGGACTAAACTCCAGAAGGAGTTATGCCTCGCTCCGAACAGAT
CTACCCCGTCCGGAAAAGTCGGACAAGGAGCAGACCAAGCTCCAGGAGCGACTGCTGAAGAA
15 GCTGGGTTCGAACGCCATCCGTTACGTTAACATCTGCCGAATGCTCCGTCTCGGTACG
CTGCAGCAGGGCGAAGATGATAATGGAGACCCGTGCGGTGTCGTACTACGTGAAGATCTTG
CCGGTGAGTCGGAAACCGATCGTACGCACCGTGCAGCACCGTTACGCTCGGCATACGCAAGAT
CCAGTCGCACCGACCAAGCAGGCCAGCAGCCGTGCACGCTGGTGCAGGACTTATGCTA
AGCCCCGGAGAGCTGGAGGTCACACTAGACAAGCAGCTGTACCTGCACGGGGAGCGA
20 ATAGGCGTCAACATCTGCATCCGCAACAACCTCGAACAAAATGGTCAAGAAGATTAAAGGCCATGG
TCCAGCAGGGTGTGGATGTGGTGTGTTCCAGAATGGTAGCTACCGAACACAGTGGCATCGCT
GGAGACTAGCGAGGGTTGCCAATTAGCCCGTCCAGTCTGCAGAAGGTAATGTACCTCACG
CCGCTGCTGTCCTCGAACAGCAGCGACGTGGCATGCCCTGGACGGTCAGATCAAGCGTCAGG
25 ATCAGTGTGTTGGCCTCGACAAACCCCTTGGCTCAACCGGATCAGCGAGATGCTTCCGGTTAT
CATATCGTATGCCGTAAAGGTTAAGCTTTCCTCGCGCACTCGCGGAGCTGTCGGCGGAA
CTTCCATTGCTGATGCACCCAAAGCCGGACCAAGGCTAAGGTCAACAGCAGGACAGCC
AGGCCGACGTAGAAACTTCCGACAGGATAACATCGACCGAGGATCAGTTGACTTGAATA
30 **GACGACGCAACGGTTGGAAATGCTACCTACTACCCAGGCATGGCTAACACGACGAACGAAC**
TACTACTACTAAGCATAAAAACAGGAAAAAAATGGAAAACCTAAAAAATGGATCATACAACCG
AACGCAAACGACCTACGACGATCTCACTTCCCCGTCTTTCATCCTAACGAAATAGAACGA
TGGTAGAAAAGGAAGATAAGATGGAGAGAAAGTCACGTGTATCAATGACGACGACTACCAAAA
35 CTGAAGACGTAACACATGTTCCCCAGCGAGCGGTAACGTTCTGACACCTCCGCTCGA
CAATGTACCTTTAAAAACATACAAATTAGAAGTCGTCTCACTACCTCAACCAATCCAGCCAC
TTTGGTATATACTTTCATAGAATCCTCTGAGCGCAAGGACCCATTGAAATTCAAGTGTATT
GTAACGCGACCAAATGCCTAGCTGAATGTTGAACGAGTTATGTACATCAAAAGATTGAATA
AAACAAAAAA
40

Figure 2

Anopheles gambiae arrestin 1 amino acid sequence (SEQ ID NO: 2)

5

MVYNFKVFKKCAPNGKVTLYMGKDFVDHVSGVEPIDGIVVLDDEYIRDNRKVFGQIVCSFRYGR
EEDEVMGLNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKKLGSNAIPFTFNISPAPSSVTLQQG
EDDNGDPCGVSYVVKIFAGESETDRTHRSTVTLGIRKIQFAPTKQGQQPCTLVRKDFMLSPGELE
LEVTLDKQLYLHGERIGVNICIRNNSNKMKVKIKAMVQQGVDFVLFQNGSYRNTVASLETSEGCI
10 QPGSSLQKVMYLTPLLSSNKQRRGIALDGQIKRQDQCLASTTLLAQPDQRDAFGVIISYAVKVKLFL
GALGGELSAELPFVLMHPKPGTKAKVIHADSQADVETFRQDTIDQQASVDFE

15

Figure 3a

Anopheles gambiae odorant receptor 1 genomic sequence (SEQ ID NO: 9)

5

Features:

- 1) Presumed Untranslated 5' and 3' regions are underlined.
- 2) Potential TATA box transcription initiation signal is double underlined.
- 3) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
- 10 4) Introns are tentatively assigned and are shown in lower case.

Exons are **highlighted**.

AGCTTGTTCATTATGTTGAAATCTAGCCCATTGTATAGTGCTGAACGACGAAGAACATACGAAAGTACCTCGT
CCGAACACTATCAACATTAATTATACCAAGCTAGAAGAAGATATTATAGTCAAGCCTAACATCATAGGAAACTT
15 AGCAAAACCATTAAATTACATGATGATAAGTCCCACCTTACCCAGCACAGGTTGAGAAGGACGAAAGTATCT
TTACGATAATATTACTCTAAGGTAGTTGAATAAAATAAAATTACGTGCAAGTGGTGGCATGGACATCATTC
GAAAGAATCTACTAAGTCATACACACACCCAAGACGACCGACGTAGTTCATCTAGAAAAAACGGTCAGCTCCATC
GAACACGTCAAGGACATACTGCGACATGCGTATGGTCAGTCCACTAGTGCCAACACTGGTCCAGGGCACTACCTT
CCGAAGCAGTAGAACCTAATGTATTGGAAATTATTAGGACATACTGCAACATGCATATGGCTAGTTCCGCTGGTACC
AACGATGGCACCAAGGACACTATCTGGCCCTGTAAAATCACTGTAAAATCTATACAAAAACGGTTACCCATACT
TTATCACAAAAACGGCAGGTGAGGGCTGGATTGCTTCAAAGCATTAGAAATATATAATTCAAAGTCCATAATCTCC
TTAAAAGATAGACAAAGTAGAGAACACATTTAGTGCTTTCTCGAGTTAGTGCCTCTCAAGTAAGCGTT
AATGCTCAATTGTTGATTGCTGGATGACTCTCGCTACGTGCTATAGTGGTCAATACTTCAATTAGATTTCAT
AATTAGTTCCAATTGTCACGGAAAACCCaCAAAAGAAAAAAACTGTATCTAGGGTGAATTTCGAGAACAA
ATTGGACACTTCATATGAAAAAGGACAGCTTTCAAAATGTTAAATAACACCGTGGATCCTTgttgattca
attctccaaattctgcagaataattctgcaaattttacaaaactgctcaaccaccaataattccaattaatcatctg
aacatttaaaactgataattaagatgagaattgcttcgtcatcacctaagaaatcgatttagttggataaaaagaa
caaattgaaatacaataaagtccctgaattttattcgaataacggcttgaactcatttattcaaaaacctttgaga
aattcctcgaaattggctccatagttctgctaacggccacttcaaaagcaagaactaacaaaatcataat
tatggcgaagtaactatcgttaccagtaatcgccattaaaactttctcaatttgcggctcgatccggctaaa
tacagagcagagtaacggaaagtgttacgtcaacgtcgctattgtataacgaggaacgcctccgaagggtgttgaagg
acctttcaattgaaaccaagtactgttccagttaaattggatagttataaaatgagccgttcaacgatcggg
catcatttgcattttcatctcgaggagaaatagatcgttaccgttaccgttgcatttgcggatccggcttgcatttgcgg
gaacccacggatggatgcgtacgttgcacggattcgttgcatttgcggatccggatccggatccggatccgg
35 GGCCACCGGAAGATAACGGATCAGGCAACCGGAAACCGGTACATCGCGTACCGTTGGCTTGGGATCATGTTCTA
CATCTGTAGGCTCTAACGCAAGCCCTATAACTCAAGgATGTGAAGGATAATTAAATgtgagtctctagtttagtattag
tgttccacctgtccataatctgtctttattggtagGACATCGCAAATGCATTGTCGTGCTTATGACTCAAGTCA
CGTTGATCTACAAGCTGGAAAAGTTAACTACAACATCGCACGGATTCAAGCTTGTCTGCCAAGCTAACTGCCACA
CTGTATCACCCGAAACAGCGCGAAGAATTCAAGgttaagcctgtggaaatatgactaaaaagagtgcataacaaacga
40 ctctccctccaaatgttagccccgttttacaatcgatGAGTGGAGTGTGTTGGCTGATGATCTTCTCATGTTGTGGC
TATCTTCACCATCATGTGGTTATGTCGCCAGCCTCGACAATGAACGTCTGCCGTGCCGTGGCT
CCGTGGACTATCACCATTGGACATAGTGACGGTACTGTTCTGTATCAAACCATTGGAATCGTCACTGAGCGCA
ACGTACAACCTCTGGACCGTACCATGTTCCGGCTTGTACACACATAAAATGGACAAATTGTGGGGCTTGGTAG
45 TATGGTAAAAACgttagttacggcgactacttgcctccagtaaggacagggagttgttccgttatgatattcatt
ttatcagCTTGGACATGACGTCCCTCCGAACGCCATTGGTCGCAACGGATGCCGAATGGAAAGAGATGCCAAAGC
GCATCGACCATCACTCCAAAGTGTACGGTACCGATGTACCGCTAAAGTAACCGAGTGTCTGCTTGCATCAAGGACATC
TTAAGgtacgaattggccaattaattgtgtcattttaaaagcttgcaccaactttcacagcttgcgtatgttgcatt
gcaggacatttccaaggatCTATCTTCCGCAAGTATGGGGCTGTGTAATTATCATTGCTAGTAAAGACATGCCAAGTGT
TACCGGGGGCGATGTTACGATGCCGATCTGCTGGCTGGGGCTTATTGCTAGTAAAGACATGCCAAGTGT
50 TTTCTGTTACGTAGGGAAATGAAATCTCTATACGgttagttggacacgttagaggaattaaatgttggaaagaata
tcaataccaaatagtatgttgcatttgcacggatccggatccggatccggatccggatccggatccggatccgg
TTCGATAAGCGTACCAAGCCAAGCAATGATAATTCTCTGCAAATgtgagatagcgggttgcagtacatc

5 ttaaatacgttcttatttcagGACTCTTAAAGATGTTACATCAAGGTGGAAAGTGTCTTGAAGGTTACGCTAAAT
CTTCACACATTGGCAGgtatgtattatgctgtggatttagcttgaataagctacaaacttgaaagtaattt
caatctgtttttagATTATGAAGCTATCGTACTCTATCTGGCCGTACTTCAGAGCATGGAATCAGAGTAATGGCG
CTAATATCCTAATGTTGAAATTATTTGTTAGATTATTGCATAAAGTAaTaTTAATTATACATCAAACGT
10 AAGCCCGCTaGTTTCAATTAGCCTTCCAAAATTATCAAATTGATTGATTGCAAGAGTTTCAGGAATT
TAATCTGATAGGATATCTGTTATCCAATAGAGGTGTGGAAGCGTCCCAAGCCATTGTTGATAGTTATAGCA
CCGTCGAGCAGTTGATCGCTGTGATCGCTAGGCGCACCTGATTATCTTATCTGCACCTGTTATGGCAAGGGCG
CTTTCACACGTTCACACAATATAATGCACATGTATAATGCATTCTACTTTAGCATTGTTACATATAATACC
AAAATTATGCATTTTATTCTCACGCAACGATTAGAGGATGACTTCACAAAGGTCCATCTAGGGTAGGAGGTATAC
15 AATTATACCTCTCAAAATCTCACAGCAATATGAGAAACAAAGGATACCAAGCATAACCCTTTTACTTGACAATT
TCATTTGATTATGTAATAAAGCACTGCaCGTCGACTTCCTAAAA

Figure 3a continued

15

Figure 3b

Anopheles gambiae odorant receptor 1 amino acid sequence (SEQ ID NO: 4)

5

MKKDSFFKMLNKHRWILCLWPPETDQATRNRYIAYGWALRIMFLHLYALTQALYFKDVKDIND
IANALFVLMTQVTLIYKLEKFNYNIARIQACLRKLNCTLYHPKQREEFSPVLQMSGVFWLMIFLM
FVAIFTIIMWVMSPAFDNERRLPVPAWFVVDYHSDIVYGVLFLYQTIGIVMSATYNFSTDPMFSG
10 LMLHINGQIVRLGSMVKLGHDVPPERQLVATDAEWKEMRKRIDHHSKVYGTMYAKVTECVLF
HKDILRIYLRAASMRVCNYHLYDTAATTGGDVTMADLLGCGVYLLVKTTSQVFIFCYVGNEISYTDKF
TEFVGFSNYFKFDKRTSQAMIFFLQMTLKDVHIKVGSQLKVTLNLTFLQIMKLSYSYLAVLQSM
ESEZ

15

10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500 510 520 530 540 550 560 570 580 590 600 610 620 630 640 650 660 670 680 690 700 710 720 730 740 750 760 770 780 790 800 810 820 830 840 850 860 870 880 890 900 910 920 930 940 950 960 970 980 990 1000

Figure 4a

Anopheles gambiae odorant receptor 2 genomic sequence (SEQ ID NO: 10)

5

Features:

10 1) Presumed Untranslated 5' and 3' regions are underlined.
2) Potential TATA box transcription initiation signal is double underlined.
3) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
4) Introns are tentatively assigned and are shown in lower case.
5) Exons are **highlighted**.

ttccagtaatccataataaaaaataataaagtaataatagtaataatagtaataattccagtaactgttagtaatac
acaataatctctaagaattaaaattgcatttgtatgaaatatgtttagttcgatgttcagaaaaacttaaa
aatgcctcagcattaaacagtttagttcagggcatttagtttagatattttatgtttagtattttaaagcatttgc
ttcattactacaaaaaagcaaatttatgagtgaattacttcagttcttaaaccgcctatgttatgcattacat
aacaatagctcttttattgcattttccttagtaatcttaaaccatctcttcccttgcagATTAAA
GTCGGCAACGTGTACCCGATGACGGTGGAAATGTTCAAAAATTGCTCAACGTGTCTACTCCTATTTCACACTGCT
GCGCCGACTGTACAACTAACCTAACCGTAAACAAACAAAAATCCCCTCATCACTATGCAAAGACAGCAAGCAGCC
GATCATCAAACACCATAGCAGCCACAAAGTTACAGCCGTTATCCCACGGGATTGGTGGAAAGTTATTGCACTG
AAGCTTTCAACCAATTTCATGGAGGTTCCCTCTCAACCAACCCATTGAAGCGAATAAAAGTATCAGCAACCAAG
GCGACGGTAAAAAACGCTGCATTATTGTGCTGCTTCAGCATTCCAGCGAATGACTCTTAAACTTTCCATTCAAA
AGTCGCGATGCTCACGATACGGAGCGGTGTTGATCCGCGAGTCACTCGCAAGCCGGTATGTTGCCAGT
GGAAATGCACAGATCGACACAGCGATAGATAATCGTTGTCGCTAAATGGGAGGGAAAAAGTAAGCTGCCAGCT
ACTTCATTCATGTTAATTGAAACTCAAGCCAACGAACATGCAGAACCCGGTGGTGTGTCTCCGCTCCGGGA
AAGGTCTCTGCTCCGGGCATGGATTCTTCCCCCTCCGGTGGTGGGGTATTGTTAGGTTTTATTTACAAA
TTCATATCCTCCGCTCCGCATCAGCCGACCCGGTGGTGCAGACAGATGTGCGGCGGGCAACAAAACATATGC
ACGAACATGGCCAACAAACACAGCTTCTATCTCATCTCTGTGTCGCACTGTCTCGCTTCCGCTGCGTTGCTGTA
GTACTATCATTGTTAGTCCACGGTTACTTCTAATTCCATTGCACCAACGCAAAAGGCTCATCCTTGCTCGTT
CCGGTTGCAACTCGACAAGCGCATGGTGGGATACGAACAAAAACCAACTACTCCACCCACTACTACTACTG
CCACCAACACTAACAAACACTACACTTGTTGGAGCTGCAGACCCACAAGCAAACACGATACAAGCTAGCTAGCT
GCTGTGCGCTCGAGTCAGCCGACGGTACAAGGTTAACCGGTACAAGCAACTCCGGACCGATCCAAAACACTCTG
ACAAGGCACGGGCCGCATCCGGCAGTACGGTGGAAAACATGGAAATGTTAATTAAAACGTAAATTGTCATCGC
TGCTACAAGTGTGACACAGGGAGAGAGAGAGACAGAGCGCGCCGATGGTGTGGTAAAGATAGATACAGGAA
AAGAGCGAGAAACATTGGTACGATTGGTGTGGTTAGCAAATTGATTCCACTGATTGAGTGCAAATTAAATGC
ATCGAAAATTGCCATTCAAGGGTAAAGTTGCTCGGACGGATCCCCGGCTGCAGGAATTGATATCAAGCTTAT
CGATACCGTGCACCTCGAGGGGGCCCGTACCCAGCTTGTCCCTTAGTGGGA

Figure 4a continued

Figure 4b

Anopheles gambiae odorant receptor 2 amino acid sequence (SEQ ID NO: 6)

5

MLIEECPIIGVNVRVWLFWSYLRRPRLSRFLVG CIPVAVLNVFQFLKLYSSWGDMSELIINGYFTV
LYFNLVLRTSFLVINRRKFETFFEGVAAEYALLEKNDDIRPVLERYTRRGRMLSIISNLWLGA FISA
CFVTYPLFVPGRGLPYGVTIPGVDVLA TPTYQVVFVLQVYLTFPACCMYIPFTSFYATCTLFALVQI
AALKQRLGRLGRHSGTMASTGHSAGTLFAELKECLKYHKQIIQYVHDLNSLVTHLCLLEFLSFGM
10 MLCALLFLLSISNQLAQ MIMIGSYIFMILSQMFAFYWHANEVLEASLGIGDAIYNGAWPDFEEPIR
KRLILIARAQPTDGGKIKVGNVYPMTLEMFQKLLNVSYFTLLRRVYN

15

Figure 5a

Anopheles gambiae odorant receptor 3 genomic sequence (SEQ ID NO: 11)

5

Features:

- 1) Presumed Untranslated 5' and 3' regions are underlined.
- 2) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
- 3) Introns are tentatively assigned and are shown in lower case.
- 10 4) Exons are **highlighted**.

AAGCAGAACACATCAAGAAGCAATTAGGTGTACGTACAGTAGTCGCAGGGAGGAATAAATAGAGGC
TTCTGAGCGGCTTCGTCTCATTAGTTCCTCGGAACTCCCCAAGACAAACGCACGATGGTACTGCCAAAATTAAAGG
ATGAAACAGCAGTGATGCCGTTCTGCTGCAAATCAAACATTGCCGGACTGTGGGTGACCGTCCACGGTAC
CGTTTATCTATCTTTCTACTTCTGCCGATCGTGGTTCTACCCAAAGTGCTTCGGTTATCCAGATCCGA
CGTTCCGCTACCCGGACCGGCAGTGCATCAGCTGCAGATTGGCAGCTAGGTGTATGTTCCTTCAC
GGCACACTACCGAGCATTGGTGCATCAGCTGCAGATTGGCAGCTAGGTGTATGTTCCTTCAC
caaacccTcgcaacatccTcgtaacactgctacatttcagTCCTCCAAGACCTACCCAAGAGGCTGGGAGACTA
CTGATCTCAGTGAACCCGACGGTCCATCGGTTCCAAAATTTACTGCTGGTGTCACTTTCCATCCAACGTTCT
TGGTTCATGGCCGTTGGACCGATTCCGCTACTTTGCTGGCAAACGCACGGTTCT
GGCACTCCGAGGAAGGCGTTACTTTOTGAACATTCGACTTCGATCGGCCACTAACGTTTTTATGTTCCAATTATG
TGGCCCACGATTAACGCTGGGTTTACCGGTGGCACAAAGGCTGGAAAAACGGCGGA
GGCACTCCGAGGAAGGCGTTACTTCCGAAATCCACTGTAGGGAGGTAGGCCAAAGACCGGCGGA
ACGAGATTTCCATCGCATCAGCCGTACTCAAGtaagtaaaatttgaaagtttgcaggaaatttgag
tgtgtcgacccgtgcacatcctagTGGTGTCCTGTGGAGGACATTCCGTGGCTATTTTTCCTGGACT
ATTCAGTTACAATGATGTGGTGCAGTCTCATCCTCTACCGGTTGGGTTCATTTCCA
gccttattcaatccattttgtgaaacgtgaaatttccccccagGGGTCAGTCGACGGTAGGGATGTATCGTCCA
ATCATTTTCGTACGGTGGAAAACTTTACGGCTACTTCGGAACAGTCAACCGGAGGTGTGGgtacc
ctttggatgaagctttcaaaagttcaaaatttgttcgattttccccttccactagAGGTAGGGGTG
CCTGGCCATTTTACGATGGGAGGTACAACTTTTTCCACTTCCACTTCCAACGA
TCCCAAAAAACCCGTTGGCCAACGGGTTGGGTGGCAAGgtaaatttaaat
tacagttgaaaaatttctgaaatgcatttacttgcttactttgttgttccagAGGTCAAAGAGTTGGCTATTTCCA
TTTACGTACTTCAACCGAGGCTTTAGGAGTGTGTTTCCACCCGTGGAATGGCTTTCGUACTGTTTCTGT
TTTGTGGACGCACGCAGCACGGGGAGGAGGCACGCCCCCGTUACGCUGTUACGCCCCC
ACGCTUAGGAGGCAGGAGGCAAATTTTGTGTTTTATTACGCUGTUACGAAATGGCCAAAG
TTTCTACGTTTTTAGGGTUTTTACTTGTUACGCUGTUACGCTTTTTCCUCAAAAACACAAAAACGTACCGACG
TATTTCTUTTTCTUGTUAGGATUAGGAGGCAGGAGGCAAACGCUGTUACGAAAGGGGGACCCCACG
GGGAAAAACCGGGAGCAAACGAGGCAGGAGGCAAACGCUGTUACGAAACAACGGTUCCGGA
ACGAG

40

Figure 5b

Anopheles gambiae odorant receptor 3 amino acid sequence (SEQ ID NO: 8)

5

MPSERLRLITSFGTPQDKRTMVLPLKDETAVMPFLLQIQTIAGLWGDRSQRYRFYLIFSYFCAMV
VLPKVLFGYPDLEVAVRGTAELMFESNAFFGMLMFSFQRDNYERLVHQLQDLAALVLQDLPTEL
GEYLISVNRRVDRFSKIYCCCHFSMATFFWFMPVWTTYSAYFAVRNSTEPVEHVLHLEEELYFLN
10 IRTSMAHYTFYVAIMWPTIYTLGFTGGTKLLTIFSNVKYCSAMLKLVALRIHCLARVAQDRAEKEL
NEIISMHQRLNCFVLLLETTFRWVFFVQFIQCTMIWCSLILYIAVTGFSSSTVANVCVQIILVTVETY
GYGYFGTDLTTEVLWSYGVVALAIYDSEWYKFSISMRRKLRLLLQRSQKPLGVTAGKFRFVNVAQF
GKMLKMSYSFYVVLKEQF

15

Figure 6a

Anopheles gambiae odorant receptor 4 genomic sequence (SEQ ID NO: 12)

5

Features:

- 1) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
- 2) Introns are tentatively assigned and are shown in lower case.

10 GGGGAACCTCCCCACCCGACCAGACGACGGAAAGCTAACGATGTGCAATTGAATAGTCATTAGT
AGCGTTTTGCTCGCAAACGAACTAACCTTGTACTTTAAGTTCACTACGGTGAGGACAAAAAA
TCAATAAAATTAAATCGAGACCGTTGATGAGCAAAAGAAAAAAAATATTTACTGATTTCATTT
CGTTCCATCGACTACATAATTATGCCACATTATTATAAGTTGTATCATTAA
AACAAACACAAAAATGCATCCTTCGAATATTAGTCAGGTTGTATCAACAATGAAGTTGAACGT
15 TTCAAAAATATTCCCTCCCCGGACACGGTCTTATCCTCGTCTAAGGCTTTGCATATCGTGGGC
ATGAATGGGGCAGGATTCCGGTCGCGAATTGAGTTGGTGGCATTTCCTGTTCTATTAAATCTT
TCTTGTAATACCGCCACTAACGGCGGGTACACCGATGGTACCCAGCGTGTACGCACCAGTGT
GAATTCCCTGTTAATTGCAATATTACGGCGGCAGTATGTTCTTGCCTACGATGTGGCCACTTT
CCAAGCGTTCATCCAGGAACGTAAAGAGCCTTCGGTTGGtaatatttaattaattaaaattgcatt
20 catcattgttctttcagTATGCTCACATTCTGACTACAGACTAAAGTATAAGCTGACCCGGTTCAACCGTC
GAGCGGATATTATGCCAAAGTGCAAACGACCTGCATGGTGCTGTAACGCTTTCTACTGGAT
TGCACCGATACTCCATCTGTGCGCACTACTACAGGTCGACCAATTCCACCGAACCCGTGCGG
TTTGTGCAACATTAGAGGTGAAGTTCTATTGGCTCGAGAATCGCACCTCAGTCGAGGACTACAT
AACCTCGTGTGATCATGCTACCCGTCGTGGTTATGTGTGGTACGTATGCAATTGAAGGTGA
25 TGACCATCTGCTGCAGCATGGACACTGTACACTGACACCAGGATGACTATAGAGATGGTAGA
GCAGTTGGAAAGCATGGCATCAGCGGAACGAACGCCAGCGCCATACGCAACGTGGGGCAGAT
GCACAGTGGTTACTGAAATGCATTAGGCTTTGAACACGTCAATCCGATCGATGCTGATGCTGC
AGTGGTTGACCTGCGTGTAAACTGGAGCATTCTCATCTATCTAACGAACGTGgttagttgtctt
30 gtttggaaatccaaaaacaaaaagatggctataattgaacttctattacagGGCATCTCGCTACAATCGTTACCGTGGT
GGTAATGTTTTCTGCCACTGCGGAAACTTTCTGTATTGTTACTTGGGACGCCGTTGCGA
CACAAACAGCAGCTGCTGGAGCACCGCACTCTATGCTACACGGTGGTACAACACTACCCAAATAGCCTT
TCGCAGCAGCATTAGGATGATGTTGAGACAGTCGCAAAGGCATGCACACATAACGGTGGGAAG
TTTTTCGCGTTAATTGGAAGAATTAGCAGGATTGTCAACTTATCCTACTCTGCTTACGTCGT
ACTTAAGGATGTAATAAGATGGATGTACAGTGAATGTTTTTTGGCTTGGCAACGAATGA
35 AGTTTCCGAATCTATATTAGATCTAGAATTAAATCTAGATGTCAATAATATGATCTGGCCATGA
CCGGTTCTGGTTTGGAACCAATTCTAAAACAATTGAAACTTAGGGCGAGGCATGAAATGTC
CCAAGAACCTATCCAAGTTCTGGAAACTACATATTACGAATCTATCCCATTATTGCTCGGAAC
GGTTTGGTGCTAAATATTGTCCAAATGTTGGTCTGGACCTATCCAGACAAAGATCTCAATT
40 TTCCTACCACTGGAACTGATTAATTGATGTAGGAAGTCATGGAGGTGTTCAAGGGAGAATTAAA
CACTAATGTTCCAACTCATTATTCAAGGGCAATTCTATTGCCCCACGGATTGATAC
GTATGTATTACTCCATTCTGGACTTGTCTATTCTGCTGCTGATTGGACGTGAAATGTTGA
GAAAAAGATTCTTATTGAGTGATACAGAGCCTTAAATACTCCTACGTTGTTGCTATTAA
GTATGGCCAGGCTAATCACAATCGCTACTAATGAACAGAACATCTCTTAATTAAACCTTCGAT
45 TGATAGTGTCAATGTCAATGTCGAGATAATTGAACGTGCAAACgATACCTACCTTAAACGGAGCAG
AACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTCGCGAGGGAGGAATAAAAT
AG

50

Figure 6b

Anopheles gambiae odorant receptor 4 amino acid sequence (SEQ ID NO: 14)

5

MKFELFQKYSSPDTVLSFVLRLLHIVGMNGAGFRSRIRVGGIFLYLIFLVIPPLTGGYTDGHQRVR
TSVEFLFNCNIYGGSMFFAYDVATFQAFIQELKSLVCSHYSYRLKYKLTRFNRRADIAKVQTTC
10 MGAVTLYWIAPIPSICAHYRSTNSTEPVRFVQHLEVKFYWLENRTSVEDYITFVLIMLPVVVMC
GYVCNLKVMTICCSIGHCTLYTRMTIEMVEQLESMASAERTASAIRNVGQMHSGLLKCIPLLNTSI
RSMLMLQWLTCVLNWSISLIYLTNVGISLQSVTVVVMFFLATAETFLYCLLGTRLATQQQLLEHAL
YATRWYNYPIAFRSSIRMMLRQSQRHAHITVGKFFRVNLEEFSRIVNLSYSAYVVLKDVIKMDVQ
NVSYSYFTLLRRVYN

15

Figure 7

ANOPHELES GAMBIAE

Preferred DNA Codons

Amino Acids			Preferred Codons				
Alanine	Ala	A	GCC	GCG	GCT	GCA	
Cysteine	Cys	C	TGC	TGT			
Aspartic acid	Asp	D	GAC	GAT			
Glutamic acid	Glu	E	GAG	GAA			
Phenylalanine	Phe	F	TTC	TTT			
Glycine	Gly	G	GGC	GGT	GGA	GGG	
Histidine	His	H	CAC	CAT			
Isoleucine	Ile	I	ATC	ATT	ATA		
Lysine	Lys	K	AAG	AAA			
Leucine	Leu	L	CTG	CTC	TTG	CTT	CTA
Methionine	Met	M	ATG				
Asparagine	Asn	N	AAC	AAT			
Proline	Pro	P	CCG	CCC	CCA	CCT	
Glutamine	Gln	Q	CAG	CAA			
Arginine	Arg	R	CGC	CGG	CGT	CGA	AGA
Serine	Ser	S	TCG	AGC	TCC	AGT	TCT
Threonine	Thr	T	ACG	ACC	ACT	ACA	
Valine	Val	V	GTG	GTC	GTT	GTA	
Tryptophan	Trp	W	TGG				
Tyrosine	Tyr	Y	TAC	TAT			

[http://www.kazusa.or.jp/codon/cgi-](http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Anopheles+gambiae+[gbinv])

5 [bin/showcodon.cgi?species=Anopheles+gambiae+\[gbinv\]](http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Anopheles+gambiae+[gbinv])

Figure 8

Name	SEQ ID NO	FIG. Reference
Arrestin 1 (cDNA)	SEQ ID NO: 1	Figure 1
Arrestin 1 (polypeptide)	SEQ ID NO: 2	Figure 2
Odorant Receptor 1 (cDNA)	SEQ ID NO: 3	—
Odorant Receptor 1 (polypeptide)	SEQ ID NO: 4	Figure 3b
Odorant Receptor 2 (cDNA)	SEQ ID NO: 5	—
Odorant Receptor 2 (polypeptide)	SEQ ID NO: 6	Figure 4b
Odorant Receptor 3 (cDNA)	SEQ ID NO: 7	—
Odorant Receptor 3 (polypeptide)	SEQ ID NO: 8	Figure 5b
Odorant Receptor 4 (cDNA)	SEQ ID NO: 13	—
Odorant Receptor 4 (polypeptide)	SEQ ID NO: 14	Figure 6b
Odorant Receptor 5 (cDNA)	SEQ ID NO: 15	—
Odorant Receptor 5 (polypeptide)	SEQ ID NO: 16	Figure 9b
Odorant Receptor 6 (cDNA)	SEQ ID NO: 17	—
Odorant Receptor 6 (polypeptide)	SEQ ID NO: 18	Figure 10b
Odorant Receptor 7 (cDNA)	SEQ ID NO: 19	—
Odorant Receptor 7 (polypeptide)	SEQ ID NO: 20	Figure 11b

Figure 9a

Anopheles gambiae odorant receptor 5 genomic sequence (SEQ ID NO: 21)

5

Predicted Exons: *ITALICIZED*, UNDERLINED AND **HIGHLIGHTED**.
Introns: lowercase.

45

Figure 9b

Anopheles gambiae odorant receptor 5 amino acid sequence (SEQ ID NO: 16)

5

MVLPKLSEPYAVMPLLLRLQRFVGLWGERRRYKFRLAFLSFCLLVVIPKVAFGYPDLE
TMVRGTAELIFEWNVLFGMLLFSLKLDYDDLVYRYKDISKIAFRKDVPSQMGDYLVRI

10

NHRIDRFSKIYCCSHLCLAIIFYWVAPSSSTYLAYLGARNRSVPVEHVLHLEEELYWFHTR
VSLVDYSIFTAIMLPTIFMLAYFGGLKLLTIFSNVKYCSAMLRLVAMRIQFMDRLDEREA

EKELEIIVMHQKALKCVELLEIIFRWVFLGQFIQCVMICLVLVYAVTGLSTKAANVG
VLFILLTVETYGF CYFGSDLTSEASCYSLTRAAYGSLWYRRSVSIQRKLRMVLQRAQKP
VGISAGKFCFVDIEQFGNMAKTSYSFYIVLKDF

15

Figure 10a

Anopheles gambiae odorant receptor 6 partial genomic sequence (SEQ ID NO: 22)

5 These are the predicted last three exons of another candidate *Anopheles gambiae* odorant receptor.

Predicted Exons: *ITALICIZED*, UNDERLINED AND **HIGHLIGHTED**.

Introns: lowercase.

10

aacaccatcttatcgcaaaattgtatattaccgttggaaagcggttccctcgtttctactctctctctgtctcttta
ttgatgcgtatgcgcgcgtgtataggctag **TTATGCTTACCGGATGTTGCGATCGCGCACGTGCTTT**
TCCGCATACGCCAGTGCACACTTGATGGCGGTGGTGTGACGTCGTGCGCACCGTT
TTCTGCTCGTGAGTCAGACCTTTCATTTCTGCAATATCCTGTTCTTCCCGACCCC
ACAGACGGTTAGACGGATATATGCTGGTAAAGTTGTCCCTTCATGCTGTGCTTCTG
ATCGAGCTGCTGATGCTGTGCGTACGGTGAGGATATTGTGGAATCGgtaggcaccaggc
ggtcatgagcgagtcgcgacttgaagctttgtttaaaacacatcagag **CCTTGGGGTGTGATTGATGCCGCT**
TACGGTTGCGAATGGTACCGGAAGGGTCGGTGGCGTCCATCGATCCGTGCTGCAA
TTATACACCGCAGCCAGCAGTCCGTACATGACCCATGGAAAATTGGCCATCCAA
ATGAGTACTTCAGTCAGgtgaggccattgtttcgtaatattcagtaagagtgcgtcttccttag
ATCCTGCAAGCTTCCGTCTACTTACCCCTCTGAAGACCGTCTACGGGAATAAgtaa
gcgcgagagagagagagcgtatgtcacccttggatgaatcaatagattctaatcatgaaccattgaaaaatgaatca
acatttcgtatgtcacaatattgtaccatttatacagcttaccacgaccaagcggttgcattcaggaccaaacacgttcga
caagccgcgtcacctgtggc

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

90

Figure 10b

Anopheles gambiae odorant receptor 6 partial amino acid sequence
(SEQ ID NO: 18)

5

LCLPDVAIAHVLFRIRQCTLGGGDDVCCAPFSARESDLFISCNILFLSRPHRRLDGYML
VKFVLFMLCFLIELLMLCAYGEDIVESPWGDZCRLRLMVPGRVGGVPSIRAANYTPQP
AVRHTDRMENLAHPNEYFQSDPASFLVLLYPPEDRLRE

10

Figure 11a

Anopheles gambiae odorant receptor 7 genomic sequence (SEQ ID NO: 23)

5

Features

1. Predicted Exons (7): **ALL CAPS, ITALICIZED, UNDERLINED, HIGHLIGHTED**
2. Introns (6): lowercase
- 10 3. 5' and 3' sequences: lowercase, dotted underlined

ccgcccggcaggtgacttacgcggctgacttgctggcgctgtttgtacggcaaacggctacacaagcgaatcgaattttcc
tatcacgctgcgcttaccagcgctgtggtaggcaaagaatgtcщаagttcattggcttgcgtctgtttgtgtgaacgtgt
15 gcacggttgcacgcctaaggttcggtgtgagccgagaagttgcagatcgaaaatctcttgcgtgtgtgtgtgcagtggaa
gcattgtgttagtggagaagtggaaaagaaaatgtgaaaatgcagttccagccgaccaagttacgtcggccttcgttgcgcac
gatgccgaacattcggttgcaggccagcggtcaacttctgttccggctacgtcaccggcccatactgtatcccaagggttac
tcctgggtggacgctgcATGGTGCTGATCCAGTTCTCGCCATCCTCGGCAACCTGGCGACGA
ACGGGGACCGACGTGAACGAGCTGACCGCCAACACGATCACGACCCCTGTTCTTCACCGCA
CTCGGTCACCAAGTTCATCTACTTTGCCGTCAACTCGGAGAACTTCTACCGGACGGCTC
GCCATCTGGAACCAAGACGAACACGCACCCCGTGTGCCCCAATCGGACGCCCGGTACCC
ATTCGATTGCCCTCGCCAAGATGCGGAAGCTGCTGGTGTGGTGTGGATGCCACCAACCGT
CCTGTCGGTTGTCGgtatgtgtatgtgtgtggccgttggaaagtgtttgcggcagaacccaaatctactgttacgc
ttgactgggttttttttcgggtggaggacggataaaatctgaaagaataattgagtcaacccacagggggatgcaag
25 acatcgcaggcagagagttgggttgcatttacccgcacccgaaatcttcacggcataagcttcacccgcgtgaaaaggaa
actccccatccctgttttttttcctctcgataaattactcatgcgttttttttttttttttttttttttttttttttttttttgc
cctactagCCTGGGTTACGATAACATTTCGGCGAGAGCGTCAAGACTGTGCTCGATAAG
GCAACCAACGAGACGTACACGGTGGATATAACCCGGCTGCCATCAAGTCCTGGTATC
CGTGGAAATGCAATGAGCGGACCGCGTACATTCTCTCATCTACCAAGGTACGTTG
GCAGGAATgtcctgcgcgtcacagttggcagtcagtgcggcaacacggcaaaaaatggactaaacccgtttcacaga
gccaacacattccatcagcaattgcatacctcgccggcgtggactggcaatgcagctacaacatcctgcctaaagttatgcaat
tcgagcacaatgttgcgtgttagggctttgtataatagtcgtttttgtcctctcgcttcaaaactctatcaacggaggaaa
tccatttcgctacaatgcctacagctcaagttcaaggtaatcgagcgggtggggatcaacttttattcatttgctaacgcccc
30 tcaacaaattctatgttctcaatggcaaagattactgcggcaccaatgcggcaacggaaaagaaaagcgacgattatga
agatgtccaaaccattgcggccgacgcatttatctgatgattgcggatggctttacttgtctgctacttcaggcacaaggaa
atgaaaccagcgcaggctcggttgcggcttgcggaggcttcaggcactgaggctgagttactaaatcgaacgattttacgattc
40 tggatccagtttatgtgtggcctgcattacagtggcaattataccctgatgttcattgcattttgttaagttgtctggtaaacg
cccgtaacgattatcttcaaagagattcttcaaagagattcaaaatgtgtataacaaatgctaacgaaatggaccgtactgg
agggttgcggaaagtaacgtttaaaatattcatcacaatcctctgcaactgtgtttatttattccatcatttaatagccgttg
gtggcggcagatgtgcgtgcgtccgcattccctccagcaagctgcgtgcgaaataatttattccatcatttaatagccgttg
45 cattttatttagcaaagcaatataaaaagcagetaaccatcccataaaacaaatgtgttttttttttttttttttttttttttttt
aagtaatggtttaccagtggaaagtgtccttccatcggtggacttcgcgtattctgtttatacaagtgcatacagaaaaaaa
ggacaaatcccttgcattggctaaaggccagcttcggtaccgcgttcggatgtcataaagttgtatgggttttttaacatt
acttccgccttaaccacctaattggactttcatgcttgcgttt
tcggcggcggcctcatcccaagtttgcggccaccaatattgccttcattatctgttaccctcgagcgttagggccgcggacgacttc
cggttaatgcaccgcattgcacggacggataatccgttggacggcgcgaaagcgactatgcggacggattgggttcgaccg

gtaccgcacgcataccgtaccgataccgaacaaacggtgtgcgcgaaagaatccctagcagccccactggcacgggtattgctt
ttggttctgttttcttccactggttgggtgcctggcgaaggctagctggctactttccggggccgcaatttctgcagccaaag
gcggcgtgctcggtggggccaaaagaat

5

Figure 11a continued

Figure 11b

Anopheles gambiae odorant receptor 7 amino acid sequence (SEQ ID NO: 20)

5

MVLIQFFAILGNLATNADDVNELTANTITLFFHSVTKFIYFAVNSENFYRTLAIWNQT
NTHPLFAESDARYHSIALAKMRKLLVLVMATTVLSVVAWVTITFFGESVKTVLDKATN
ETYTVIDIPRLPIKSWYPWNAMSGPAYIFSFIYQVRWRNGIMRSLMELSAASLDTYRPNSSQ

10

LFRAISAGSKSELIINEEKDPDVKDFDLSGIYSSKADWGAQFRAPSTLQTFDENGNGNP
NGLTRKQEMMVRSIAKYWVERHKHVVRLVSAIGDTYGPALLHMLTSTIKLTLLAYQA

15

TKIDGVNVYGLTVIGYLCYALAQVFLCIFGNRLIEESSSVMKAAYSCHWYDGSEEAKT
FVQIVCQQCQKAMTISGAKFTVSLDLFASVLGAVVTYFMVLVQLK